U.S. Facilities Overview

Michael Ernst
April, 2010

Responsibilities of the Tier 1 and Tier 2 Centers

The responsibilities of the U.S. ATLAS Tier 1 site at Brookhaven National Laboratory include:

- Provide reliable storage of complete sets of ESD (currently on disk plus previous version on tape), AOD, Ntuples, and TAGs on disk plus a fraction of RAW data as well as all U.S. generated RDO (Raw Data Objects) data: Monte Carlo, and Primary data. The fraction of RAW data on disk varies from site to site, but is anticipated to be roughly 10% per Tier 1 center. The fraction of ESD data on disk varies from site to site and is expected to average 20% per Tier 1 center. However, the U.S. Tier 1 center plans to hold 100% of the ESD. While we are planning to keep 100% of the AODs at the U.S. ATLAS Tier-1 center the plan is to only a fraction of them distributed over the rest of the Tier-1 cloud.
- Anticipated, but not determined yet: 100% of all dESDs are to be stored at all Tier 1 sites.
- Provide CPU for centrally managed ATLAS-wide production, primarily for reprocessing of RAW data and group analysis (dESD production)
- Provide CPU for regional and local production of large samples through PanDA

The responsibilities of the Tier 2 Cloud in the U.S. include:

- Reliable storage of RAW, ESDs, AODs, and TAGs on disk for Monte Carlo and Primary Data. The fractions of RAW and ESD formats will be trace amounts for debugging and code development. The fraction of AODs at Tier 2 sites in the U.S. is not determined: during early running, 100% or a smaller fraction, depending on the available storage resources, of AODs are expected. During long-term, stable running AODs are expected to be distributed across the U.S. Tier 2 Cloud. Individual sites may hold up to a complete AOD dataset (33% according to the computing model).
- Anticipated, but not determined yet: the hope is that multiple copies of all dESDs will be distributed across the entire U.S. Tier 2 Cloud, so that multiple sites might hold a complete replica or at least a fraction of a replica of the same data.
- Not determined yet: what fraction of D2PD/dAOD data will be available.
- 50% of CPU resources are centrally managed for Monte Carlo production and other ATLAS-wide responsibilities.
- A yet undetermined fraction of CPU resources are likely to be devoted to D2PD/dAOD and D3PD/ntuple production.

The U.S. ATLAS Facilities Integration Program

The following paragraph describes the Integration Program of work for the U.S. ATLAS Facilities, establishing a baseline set of deliverables and schedule which successfully integrates Tier1 and Tier2 fabric resources, ATLAS software and services, grid level

services, operational components, and user capabilities to complete the US ATLAS Facility. The program is organized in phases, envisioning a series of phased hardware and software integrations and deployments over a quarterly time horizon, overlapping with the reporting period of the U.S. ATLAS Operations Program. At the end of each phase a summary report, a deliverable to the U.S. ATLAS Facilities Manager, is produced which captures the progress made, experience gained and issues encountered. The Integration Program is part of the Facilities WBS (2.3.6).
As to the schedule the program is organized as a sequence of releases tied to ATLAS milestones and the ATLAS schedule. The Facility Manager and the Manager of the Integration Program (Rob Gardner, UoC) are tightly connected with ATLAS Distributed Computing operations and the WLCG.
Besides individual facility component integration (e.g. compute and storage server hardware, storage management software, batch systems) the integration of the OSG middleware is an important topic. OSG provides the core middleware stack needed by PanDA, DQ2, VO security/authorization infrastructure, and individual user access to distributed U.S. ATLAS resources. OSG also provides centralized services for accounting, site, service verification and logging. The facilities with several instances at the Tier-1 center and at Tier-2s are participating in the OSG Integration Test Bed which provides a pre-release validation environment to check that ATLAS requirements are being met. The Integration Program provides the context for managing the U.S. ATLAS Facilities requirements for OSG.

Based upon the results achieved during the computing challenges and at the start of ATLAS data taking we can confirm that the tiered, grid-based, computing model is the most flexible structure currently conceivable to process, reprocess, distill, disseminate, and analyze ATLAS data. We found, however, that the Tier-2 centers may not be sufficient to reliably serve as the primary analysis engine for more than 400 U.S. physicists. As a consequence a third tier with computing and storage resources located geographically close to the researchers was defined as part of the analysis chain as an important component to buffer the U.S. ATLAS analysis system from unforeseen, future problems. Further, the enhancement of U.S. ATLAS institutions' Tier-3 capabilities is essential and is planned to be built around the short and long-term analysis strategies of each U.S. group.

An essential component of this strategy is the creation of a centralized support structure, because the considerable obstacle to creating and sustaining campus-based computing clusters is the continuing support required. In anticipation of not having access to IT professionals to install and maintain these clusters U.S. ATLAS at each institution has spent a considerable amount of effort to develop an approach for a low maintenance implementation of Tier-3 computing. While computing expertise in U.S. ATLAS was sufficient to develop ideas on a fairly high level only the combined expert knowledge and associated effort from U.S. ATLAS and OSG facilities personnel has eventually resulted in a package that is easily installable and maintainable by scientists.

Within U.S. ATLAS computing all activities associated with development, integration and operation of Tier-3 services and components are coordinated through the Integration Program.

U.S. ATLAS Tier-1 Center High Level Overview

Introduction

The production chain for ATLAS data is described below. It consists of the successive reduction of data from RAW to manageable sizes, suitable for repeated analysis. This reduction is performed at increasing detail through an international array of Tiered computing centers. There are ten national computing hubs called Tier 1 centers in the U.S., Canada, Germany, the United Kingdom, France, Italy, Scandinavia, the Netherlands, Spain, and Taiwan. Associated with each Tier 1 center is a set of Tier 2 and Tier 3 clusters. Tier 1 and Tier 2 centers provide ATLAS-obligated resources and the tasks which they perform are defined by ATLAS computing and physics management. For example, Tier 1 centers have responsibilities for production tasks which are ATLAS-wide, in addition to reprocessing and other responsibilities. Tier 2 centers are required to provide a minimum of 50% of their resources to ATLAS-directed simulation effort and the other 50% to computing needs associated with analysis.

As described in the Computing Model (CM) and the Computing Design Report (CDR) like all Tier-1 centers the U.S. ATLAS Tier-1 center provides a full-service computer center for the collaboration. According to its designated share the center provides a fraction of 23% of the raw and simulation data storage for the experiment. The center also provides the necessary resources to perform reprocessing and analysis of the data. In accepting to host a fraction of the copy of the raw data the U.S. ATLAS Tier-1 center also accepts to provide access to this data for the entire collaboration and to provide the computing resources for the reprocessing. As described in the CM it is not expected that access to all of the hosted raw data is available with low latency but a fraction of at least 10% is kept available on disk storage for calibration and algorithm development.

However, access to ESD, AOD, DPD and TAG datasets should always be possible with short latency (on 'disk'), at least for the most recent version of processing while previous version(s) will be available with higher latency (on 'tape'). In accepting data from Tier-2 centers a Tier-1 accepts to store them in a permanent and safe way and to provide access to it in agreement with current ATLAS policy. This is true for both simulated and derived data.

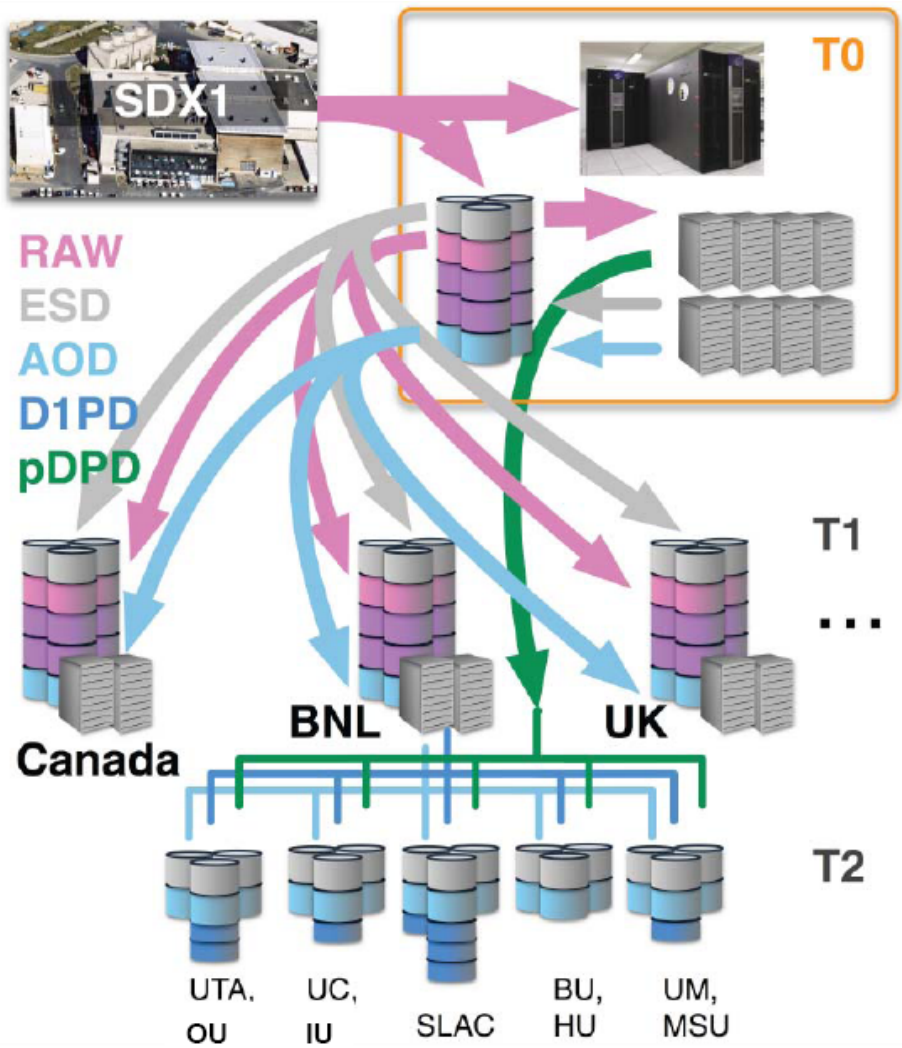


Figure 1: The data flow from the detector (SDX1) through the CERN-based Tier-0; the set of Tier-1 centers and through the U.S. ATLAS Tier-1 center at BNL to the U.S. Tier-2s

In order to meet these goals the following set of services are provided:

- Mass storage data archive
- Infrastructure for site security and access restrictions.
- Prioritization of access to data and processing resources, in agreement with ATLAS policy.
- Accounting information both for processing and data storage.
- Database services to allow replication and caching of database information for calibration data, parameter set of data, data set bookkeeping data, etc. in agreement with the ATLAS DDM strategy.
- Publication of the necessary information to be used by the Grid services.

The services in the area of processing and storage hardware, grid middleware, network and services related to the general computing environment (e.g. LDAP) are provided in terms of the level of availability and response time in case of problems according to the metrics as defined in the WLCG MoU. Resource utilization is organized through the ATLAS Production System in agreement with the U.S. ATLAS Production Coordinator and the general ATLAS production policy.

The following parameters define the minimum levels of service for the Tier-1 center. They are reviewed by the operational boards of the WLCG Collaboration.

Service	*Maximum delay in responding to operational problems*			*Average availability[5] measured on an annual basis*	
	Service interruption	Degradation of the capacity of the service by more than 50%	Degradation of the capacity of the service by more than 20%	During accelerator operation	At all other times
Acceptance of data from the Tier-0 Centre during accelerator operation	12 hours	12 hours	24 hours	99%	n/ a
Networking service to the Tier-0 Centre during accelerator operation	12 hours	24 hours	48 hours	98%	n/ a
Data-intensive analysis services, including networking to Tier-0, Tier-1 Centres outwith accelerator operation	24 hours	48 hours	48 hours	n/ a	98%
All other services[6] – prime service hours[7]	2 hour	2 hour	4 hours	98%	98%
All other services[6] – outwith prime service hours[7]	24 hours	48 hours	48 hours	97%	97%

The response times in the above table refer only to the maximum delay before action is taken to repair the problem. The mean time to repair is also a very important factor that is only covered in this table indirectly through the availability targets. All of these parameters will require an adequate level of staffing of the services, including on-call coverage outside of prime shift.

The following parameters define the minimum levels of service for a Tier-2 center. They are reviewed by the operational boards of the WLCG Collaboration.

Service	*Maximum delay in responding to operational problems*		*Average availability[5] measured on an annual basis*
	Prime time	*Other periods*	
End-user analysis facility	2 hours	72 hours	95%
Other services[6]	12 hours	72 hours	95%

As to the steady state data distribution of a number of operations automatically flow from the Tier-0 center at CERN, pushing data to the Tier 1's. The ESD, AOD, and TAGs are Tier-0 responsibilities and are cached at the Tier 1 center (along with RAW). The D1PD/dESD format is subsequently created at the Tier 1 from the ESDs. Table 2 lists the operations, including the point of origin, destination, actual computational responsibility, as well as the group responsible for the operation.

data in:	data out:	from:	to:	by:	trans:	who:
ESD	ESD	T0	T1	T0	C	
AOD	AOD	T0	T1	T0	C	all groups
AOD	AOD	T1	T2	T1	C	all groups
AOD	D1	T1	T1,T2	T1	SK, SL, TH	all groups
ESD	pDPD	T0,T1	T2,T3	T0,T1	SK, SL, TH, AU	all groups

Table 2: Steady state data distribution use cases

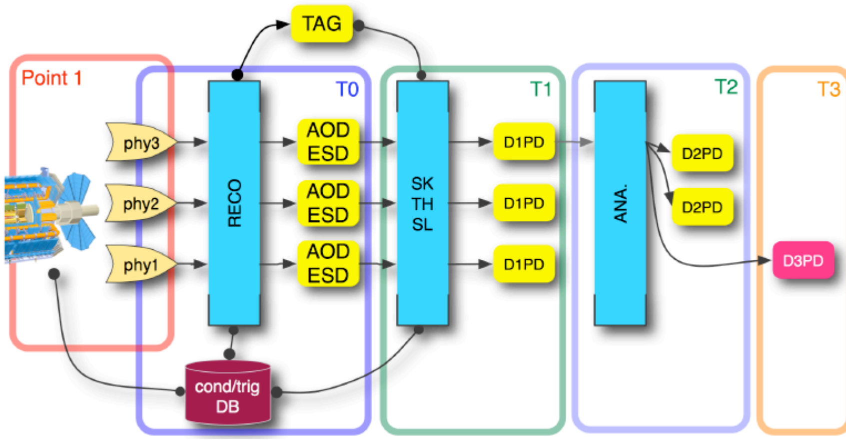


Figure 2: Production stages from HLT through the D3PD

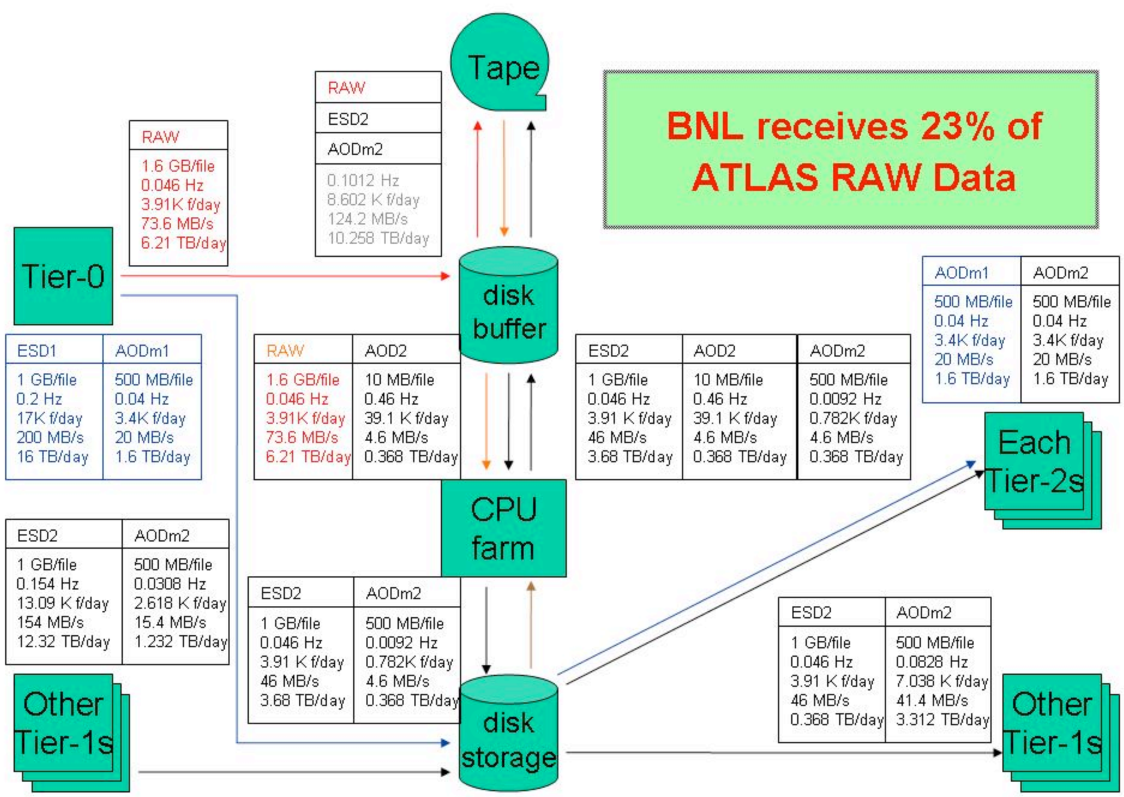
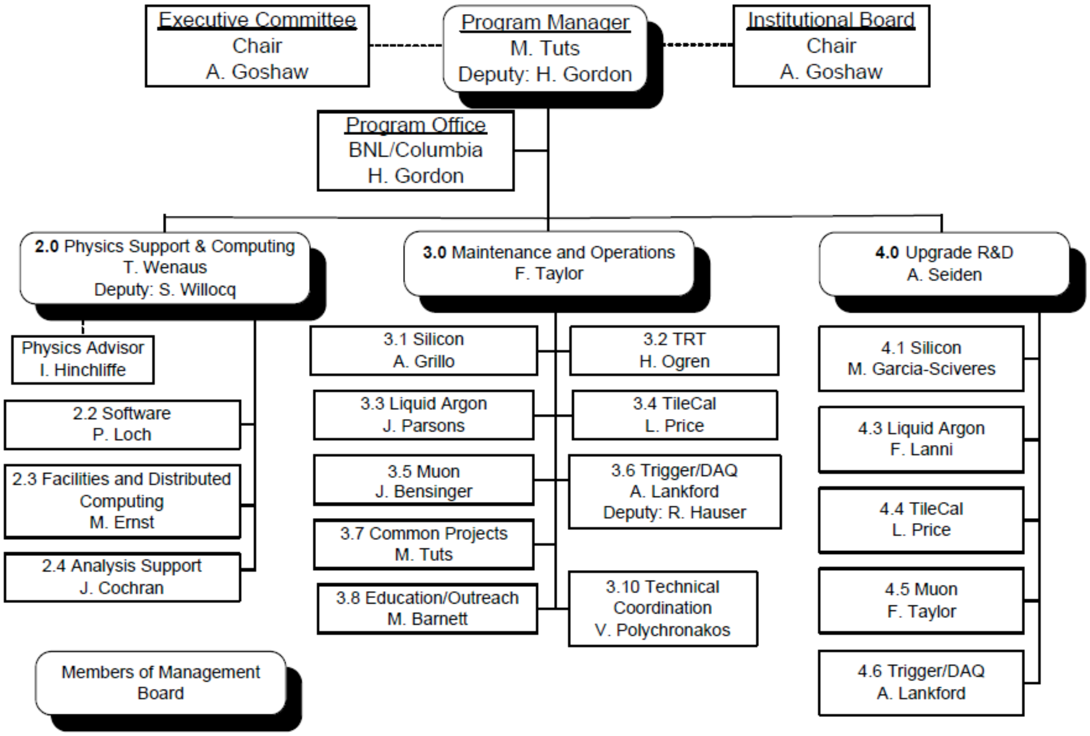


Figure 3: U.S. ATLAS Tier-1 data flow in 2009 (plus simulation and analysis data flow)

The Facilities in the U.S. ATLAS Organization (WBS 2.3)

The U.S. ATLAS Operations Program Organization as of February 2010



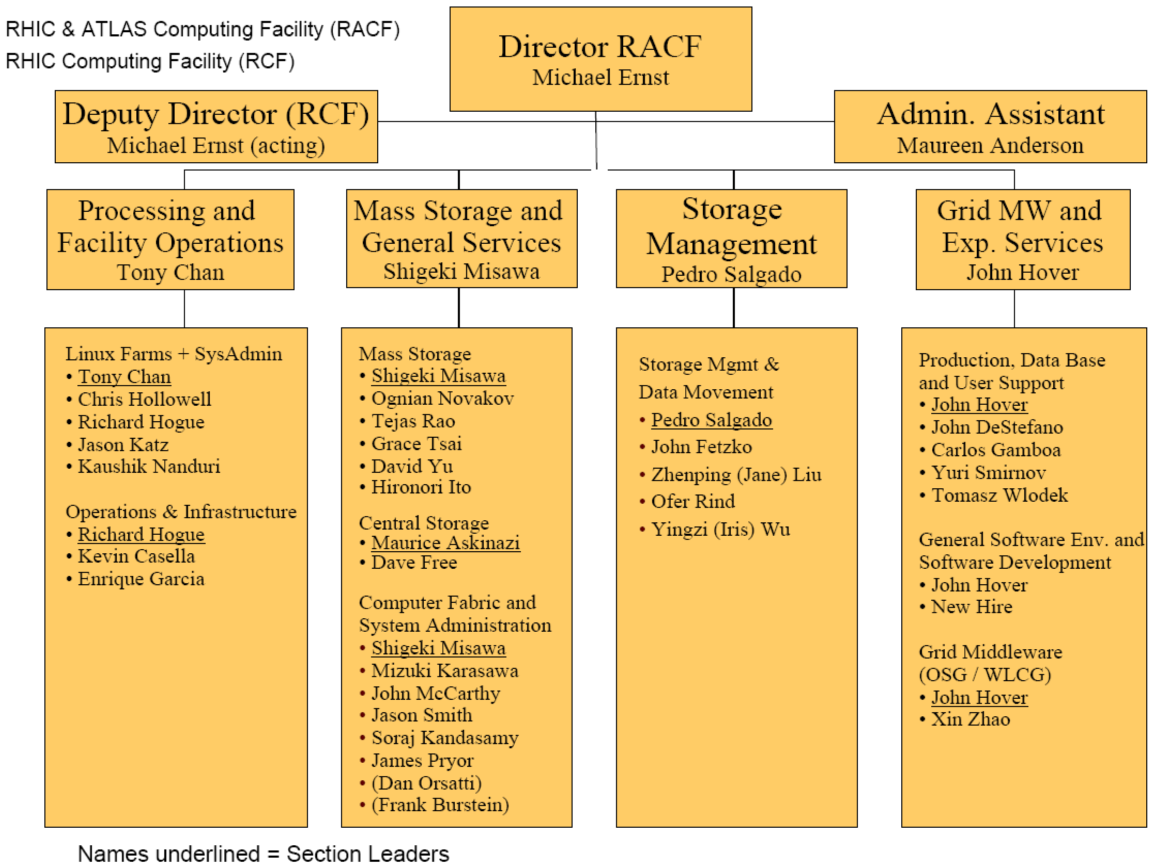
WBS 2.3 Computing Facilities in the U.S. ATLAS Reporting System

2.3. Computing Facilities
2.3.1 Tier-1 Facility
2.3.1.1 Management/Administration
2.3.1.2 Tier-1 Fabric Infrastructure
2.3.1.3 Tier-1 Linux Systems
2.3.1.4 Tier-1 Storage Systems
2.3.1.5 Tier-1 Wide Area Services
2.3.1.6 Tier-1 Operations

The Tier-1 Facility

The RHIC and ATLAS Computing Facility (RACF) is administered in the BNL Physics Department. The RACF is a dedicated computing facility, located at BNL, with an architecture based on farms and arrays of cost-effective commodity components configured specifically to serve the needs of the RHIC experiments and the ATLAS experiment at CERN. The RACF is accessible to scientists at remote collaborating institutions via high quality network connections. It provides the computing resources needed to record the data produced by the RHIC detectors and the ATLAS data that needs to be stored at the Tier-1 center, archive the data and perform event reconstruction and analysis functions.

Authorized staff is 20 FTE for the RHIC and 20 FTE for the ATLAS computing facility with effort organized according to services rather than programs. It follows the concept of a layered model for providing support for different science programs, sharing as much as possible, and by leveraging, consolidating, and focusing on robust solutions to drive down the risk, and the cost of operations. Since there is a lot of overlap across the services needed in support of RHIC and ATLAS there are substantial savings due to synergy. According to calculations made both facilities would need ~6 FTE more if they were run as separate entities.



The Tier-1 Facility Subsystems

2.3.1.2 Tier-1 Fabric Infrastructure

The primary mission of the group that is concerned with deliverables in the area of Tier-1 fabric infrastructure is to handle hardware provisioning, OS software support, AFS and NFS network file system support, basic web services, general security infrastructure, local area network, user account management, general user support utilities (subversion, Request Tracker), monitoring (system, services and applications, Nagios based) and user database (MySQL) (responsibility shared with the Grid Middleware and Services Group).

2.3.1.3 Tier-1 Linux Systems

The Processing and Facilities group has two major roles:

1) managing and operating the 6500-core farm cluster as part of the RHIC Tier0 and US-ATLAS Tier1 computing centers, and
2) overseeing and managing the physical infrastructure for the entire RACF complex.

Major responsibilities include lifecycle management of the compute server hardware (evaluation, acquisition, installation, maintenance and retirement), operation of the cluster management software stack (Linux OS, Condor batch, Xen virtualization, security, etc), organizing the usage of floor space, cooling and power in the data center, planning for capacity increases in computing resources and infrastructure needs to meet programmatic requirements, and real-time monitoring of the usage and performance of computing and infrastructure resources.
There are currently ~2000 compute servers maintained by the group. Taking advantage of scalable management solutions the group has developed they would need one more FTE if the number of servers were to be doubled to 4000.

2.3.1.4 The Tier-1 Storage Systems

The storage services at the U.S. ATLAS Tier-1 center can be subdivided in 2 areas

1) The disk based storage area with data accesses services as they are used by ATLAS applications with transparent migration and retrieval capabilities for archived data
2) The magnetic tape based archival storage (HPSS)

The disk based storage area provides a usable capacity of currently 3.8 PB and is supposed to grow, according to requirements presented by ATLAS management, to 6.0 PB by the June of 2010. As to the technology used the majority of the space is furnished by high-density storage servers from Sun Microsystems (recently acquired by Oracle) that offer in the configuration used by the facility a usable space of 33 TB per box. The chassis comes with 48 disk drives, CPU (8 cores) and a 10 GE network interface. This setup is augmented by 2 Nexsan storage arrays (connected via Fiber Channel to the Sun storage servers) per Sun server providing ~100 TB/160 TB of usable space per manageable unit.

The former approach the facility has taken in terms of using “disk-heavy” worker nodes with up to 5 TB of usable disk space per unit has been phased out in late 2007. Primary reasons were the large number of manageable units and stability problems with static and transient failures. 1 PB of disk space provided on aging worker nodes was retired over the course of the summer of 2009.

As more storage capacity needs to be added shortly and given the uncertainties as to the future of Sun as a hardware company intensive market research is being conducted. Companies under consideration include Data Direct Networks (DDN) and a few others. DDN has agreed to deliver a 2 PB system for evaluation purposes. The unit will be, following initial tests, integrated into production for a period of several weeks before a purchasing decision will be made.

At the facility a storage management solution based on dCache, a software system developed by primarily DESY and Fermilab, is used on top of individual storage servers. dCache, used at seven out of ten ATLAS Tier-1 centers, offers a scalable solution that is capable of managing a large number of heterogeneous storage servers while maintaining a homogeneous file system view. It offers, besides full file copies in and out of the system, POSIX-like data access capabilities. By utilizing integrated standardized wide area network interfaces dCache is fully integrated into the ATLAS Distributed Data Management system (DDM) and the ATLAS production system.

While the performance and stability of the dCache system has significantly improved over time, recent ATLAS exercises have unveiled some serious concerns as to primarily two dCache components: the namespace manager (PNFS) and the Storage Resource Manager (SRM) interface. The facility is closely tracking developments in both areas and has implemented a data access method that is designed to largely reduce the stress on PNFS. We are also evaluating Chimera, the successor to PNFS. Depending on the results a possible migration may take place toward the end of 2010 or later.

Members from the Storage Management Group are in close contact with experts from other sites regarding alternative storage management solutions. Candidates of interest include STORM/GPFS and BestMan/Lustre, but also commercial solutions like BlueArc.

In order to maintain an operational instance of the US ATLAS dCache at BNL, a person is assigned as the ATLAS Storage Manager On-Duty (SMOD) and another one as her backup. The person assigned acts on a weekly basis in this role.

The storage manager on-duty is pro-actively monitoring the storage services and she is responsible to respond to RT Tickets assigned to the Storage Management Group.

The Mass Storage Group is running a single HPSS instance to support the archival needs of the RHIC and the ATLAS programs. Out of the total data volume of ~10 PB currently managed by HPSS ATLAS has a fraction of 3.0 PB. Common to both research programs is the HPSS core server, while the data mover infrastructure is dedicated and sized according to the needs of the respective programs. The group maintains the entire server infrastructure (core server, mover) including the mover network interconnect with dedicated Ethernet switches, the high-performance HPSS disk cache that acts as an intermediate adaptation layer between the application buffer (dCache read/write pools) and the tape drives, the tape libraries (through a maintenance contract), the HPSS software installation and configuration, and the system that is used to manage the incoming retrieval requests to optimize the tape I/O. The group manages all operational aspects associated with the above.

2.3.1.5 Tier-1 Wide Area Services

The group is primarily concerned with support of the grid middleware services as they are required by the ATLAS production system (OSG middleware installation, configuration and maintenance incl. an instance of the OSG integration test bed (ITB), Condor-G based grid job submission infrastructure for the US ATLAS VO, VO authentication system (GUMS), certificate management, Computing Elements, the FTS server/service, the LHC file catalog used to keep track of the file inventory in the US, smart switch based resilience mechanisms, cyber security support for the US ATLAS VO, grid-based reliability measurement (RSV) and grid-based accounting (Gratia)) and to support group and user analysis. The group maintains and operates the ATLAS-specific site service infrastructure for the ATLAS distributed data management (DDM/DQ2) that is used for data replication within the US cloud and in between the US ATLAS Tier-1 center and other ATLAS Tier-1s. They also spend some effort on the optimization of long-haul data transfers, in particular between the Tier-1 center and the Tier-2's (part of the Facility Integration Program). They maintain and further develop service oriented monitoring capabilities to detect facility related component failures. Database support in terms of Oracle (used for conditions and TAG data, FTS, LFC and in the future for Chimera) as well as MySQL, as it is needed in the ATLAS Computing Facility, is entirely located in this group.

The group has also taken major responsibility for areas, though associated with facility services that are not solely Tier-1 tasks. Examples include development, implementation and operation of the central PanDA servers since the decision was taken by ATLAS to use PanDA as the global production system (from September 2007, service migration to CERN was completed in early May, 2009). Another example is work on optimized and scalable access to conditions data using FronTier, a caching mechanism based on squid (web caching technology), developed by CDF and successfully adopted by CMS. Following extensive testing FronTier was adopted ATLAS-wide and was globally deployed under the co-leadership of a member from this group.

2.3.1.6 Operations

Computer center operations comprises primarily effort spent on development, planning, implementation and operations of computer room infrastructure (space, power and cooling), and effort associated with operating the services (includes magnetic tape cartridge handling etc).

The Tier-2 Centers

The Tier-2 centers in the distributed computing model bring with them a number of other advantages, apart from that of gathering and managing resources worldwide to meet the aggregate computing needs. First by balancing on-demand local use and centrally coordinated production use, the experiment can ensure that individual scientists and small workgroups have the means at their disposal to develop new software and new lines of analysis efficiently. In many cases, hosting frequently accessed data close to the users leads to more rapid turnaround, as a result of higher throughput achievable over local area and/or shorter wide area network links.

Second, placing Tier-2 centers at universities and research laboratories combines the research and education mission, making students and young scientists part of the ongoing process of exploration and discovery, including many who will rarely be able to visit the central laboratory.

Third, by having appropriately scaled centers in the U.S. region, the scope of the analysis will be increased, proportional to the increased number of intellectual focal points for student-faculty interaction and mentoring. The distributed nature of the Tier-2 centers and their flexible balance between on-demand use by local and regional groups and centrally managed/organized use also leads to new modes of partnerships between the universities and laboratories, with greater continuity and less reliance on people's ability to travel. This leads to more effective and creative collaborative work among members of small working groups.

The development and ramp-up of resources and operation of the Tier-2 centers is coordinated through the facilities Integration Program (described earlier in this note), which has created working groups to address specific topics including Distributed Data Management (DDM, includes configuration and resource allocation of storage space according to ATLAS policies), optimization of network throughput between the Tier-1 and Tier-2 centers and technical optimization of analysis capabilities at the Tier-2 centers. Apart from using technical forums, organizational and technical issues are addressed in a weekly meeting chaired by the leader of the Integration Program, Robert Gardner from University of Chicago and Michael Ernst from BNL. A comprehensive agenda is always submitted prior to the meeting and minutes are available right after. Workshops with typically 40 participants attending in person are held on a quarterly basis. Good and steady participation from the sites is observed for more than three years. Site administrators from all nine locations hosting Tier-2 components, the Tier-3 coordinator and technical experts responsible for facility and production services are advised by the facility manager about upcoming ATLAS processing and data replication plans, and they review the ongoing production and analysis activities with respect to the performance of the U.S. ATLAS facility resources. If necessary, action items are created, documented and tracked by the Integration Program.

As to human effort at the nine institutions hosting Tier-2 resources and services there is about 1 FTE per site supported by the U.S. ATLAS Operations Program and on average 0.5 to 1 FTE contributed by each site in addition. With this amount of effort the 5 Tier-2 centers are capable of managing and maintaining about 10000 job slots and 5 PB of disk storage.

U.S. ATLAS Operations

Recognizing that production operations is increasingly gaining importance as we were getting prepared for LHC startup, concerted activities in this area managed under facilities started in June 2008. With the creation of the ATLAS Distributed Computing (ADC) organization at CERN in early 2008 with strong focus on worldwide computing operations, the development in U.S. ATLAS is intended to match the central effort with a seamlessly integrated regional one. Kaushik De from UTA was appointed U.S. ATLAS

Operations Coordinator, with Armen Vartapetian as deputy for production related operations issues and Nurcan Ozturk as deputy for all facility related user analysis support issues. In addition to the two deputies there is the shift captain organizing and overseeing computing shifts in the U.S., a role filled by Yuri Smirnov from BNL. Yuri also supports production activities in terms of task validation and complex error analysis.

As to the Facility WBS, computing operations is organized under WBS 2.3.5.

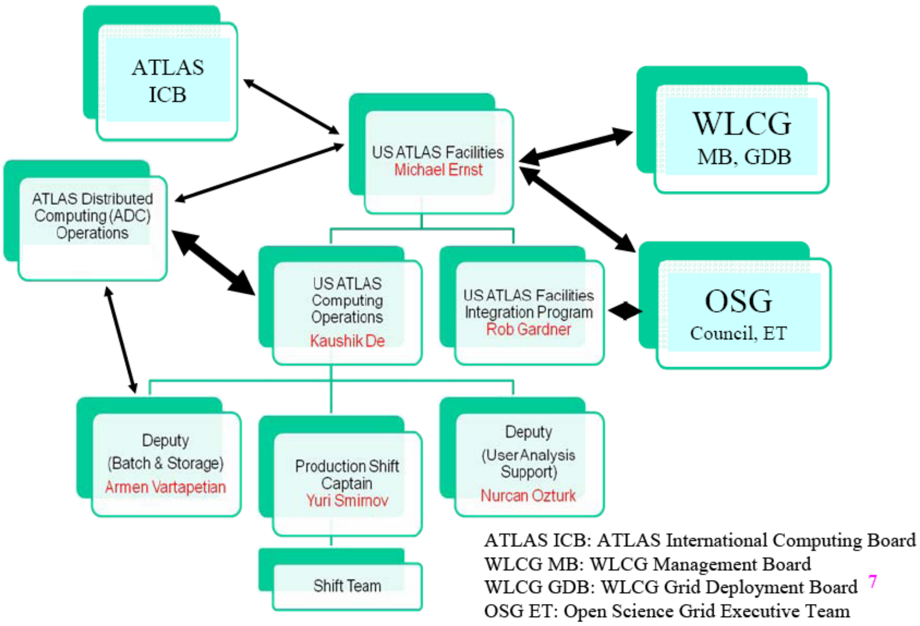


Figure x: U.S. ATLAS Computing Operations organization

Daily operation of computing and storage resources as they are available to ATLAS in the U.S. comprises

- Data production shifts (MC and reprocessing)
- User analysis support shifts
- Distributed service monitoring – grid operations and data management
- Site/resource utilization
- Overall quality of service monitoring and assurance

Planning and scheduling is another task associated with computing operations. Main activities include coordination with ADC and resource allocation (CPU and storage) at the Tier-1 center and the Tier-2s in the U.S. Besides the latter the coordinator and his deputies interact with experts developing the ATLAS software framework and the system for production and distributed analysis and, in support of operations, they oversee the development of monitoring tools.
The operations team in the U.S. interacts strongly with the ADC team organizing ATLAS computing operations centrally. The team participates in the weekly

- ADC operations meetings
- ADC operations shifts meetings
- Distributed Analysis shift team meetings
- ADC development meetings

U.S.ATLAS Operations Program Facility Personnel Tier-1 Facility					
WBS			**Inst.**	**Class.**	**FY 2010 FTE**
2.3.1.1	***Tier-1 Administration***		BNL		0.50
	Administrative Assistant		BNL		0.50
				Effort (FTE):	**1.00** /
2.3.1.2	***Tier-1 Fabric Infrastructure***		BNL		
	Facility Networking		BNL/ITD		0.50
	User account management		BNL		0.50
	Management of storage servers		BNL		0.50
	Ops of general Comp Infrastr.		BNL		0.50
	Facility Networking		BNL/ITD		0.25
	Facility Architecture dev		BNL		0.50
	Inst & cust of facility servers		BNL		0.50
				Effort (FTE):	**3.25 /**
2.3.1.3	***Tier-1 Linux Systems***				
	Linux farm and comp room ops		BNL		0.30
	Leader Processing Services		BNL		0.50
	Dev, impl, ops Linux farm		BNL		0.50
	Compute farm operations		BNL		0.50
	Inst, cust of LRMS (Condor)		BNL		0.44
				Effort (FTE):	**2.24 /**
2.3.1.4	***Tier-1 Storage Systems***		BNL		
	Storage Management Sys ops		BNL		1.00
	Leader Storage Mgmt Group		BNL		1.00
	Impl & ops Storage Mgmt Sys		BNL		1.00
	Design & impl of Storage Sys		BNL		0.35
	Storage Management Sys ops		BNL		1.00
	Design & impl of mass storage		BNL		1.00
	Operations of mass storage H/W		BNL		0.82
				Effort (FTE):	**6.17 /**
2.3.1.5	***Tier-1 Wide Area Services***				
	Doc services, Oracle DBA		BNL		0.50
	Dev, plan, impl Oracle DB serv		BNL		1.00
	Grid-based proc services dev		BNL		0.33
	Distrib Facility Monitoring dev		BNL		0.50
	Leader Wide Area Services		BNL		1.00
	Distrib Comp Services dev&ops		BNL		1.00
				Effort (FTE):	**4.33 /**
2.3.1.6	***Tier-1 Operations***				
	Dev, plan, ops comp room infra		BNL/ITD		0.04
	Ops Computer Room infra		BNL/ITD		0.02
	Planning, impl Comp Room infra		BNL		0.20
	Inst, ops Computer Room infra		BNL/ITD		0.05
	Facility services operations		BNL		1.00
				Effort (FTE):	**1.31 /**
2.3.1	**Tier-1 Facility**			**Total Effort (FTE):**	**18.30**

U.S.ATLAS Operations Program Facility Personnel

Tier-2 Facilities

WBS		Name	Inst.	Class.	FY 2010 FTE
2.3.2	***Tier-2 Facilities***	***Michael Ernst***	BNL		
				Effort (FTE):	**/**
2.3.2.1	University of Chicago/Indiana	***Robert Gardner***	UC		
	Facility dev & ops	Aaron van Meerten	UC		1.00
	Facility dev & ops	Sarah Williams	IU		1.00
	Facility dev & ops	Nathan Yehle	UC		0.33
	Facility dev & ops	Frederick Luehring	IU		0.15
				Effort (FTE):	**2.48 /**
2.3.2.2	Boston University/Harvard	***Jim Shank***	BU		
	Facility dev & ops	Saul Youssef	BU		1.00
	Facility dev & ops	A Abaris	BU		0.50
				Effort (FTE):	**1.50 /**
2.3.2.3	UTA/Oklahoma	***Kaushik De***	UTA		
	Facility dev & ops	Patrick McGuigan	UTA		1.00
	Facility dev & ops	Victor Reece	UTA		0.50
	Facility dev & ops	Horst Severini	OU		0.00
	Facility dev & ops	Karthik Arunachalam	OU		0.50
				Effort (FTE):	**2.00 /**
2.3.2.4	Great Lakes (UM & MSU)	***Shawn McKee***	UM		
	Facility dev & ops	Robert Ball	UM		1.00
	Facility dev & ops	Ben Meekhof	UM		0.10
				Effort (FTE):	**1.10 /**
2.3.2.5	SLAC	***Richard Mount***	SLAC		
	Facility dev & ops	Wei Yang	SLAC		1.00
				Effort (FTE):	**1.00 /**
2.3.2	**Tier-2 Facilities**			**Total Effort (FTE):**	**8.08**

U.S.ATLAS Operations Program Facility Personnel

Grid Production / Facility Integration / Tier-3 Coordination

WBS		Name	Inst.	Class.	FY 2010 FTE
2.3.5	***Grid Production***	***Kaushik De***	UTA		
2.3.5.4	Dep Production Coordinator	Armen Vartapetian	UTA		1.00
	Dep Distrib Analysis Ops Coord	Nurcan Ozturk	UTA		1.00
	Production Operations	Mark Sosebee	UTA		1.00
	Production Shifts	Rupam Das	UTA		1.00
	Central Production Operations	Pavel Nevski	BNL		0.50
	Production ops	Wensheng Deng	BNL		1.00
	Production ops, US shift captain	Yuri Smirnov	BNL		1.00
				Effort (FTE):	**6.50 /**
2.3.6	***Facility Integration***	***Robert Gardner***	UC		0.40
				Effort (FTE):	**0.40 /**
2.3.7	***Tier-3 Coordination***	***Rik Yoshida***	ANL		
	Tier-3 Technical Support	Douglas Benjamin	Duke		0.50
				Effort (FTE):	**0.50 /**

U.S. ATLAS Facility Capacities

Site	CPU Installed (HS)	CPU Pled. 2010 (HS)	CPU needed in 2011 (HS)	Disk Installed (TB)	Disk Pledged for 2010 (TB)	Disk needed in 2011 (US share) (TB)
T1	31,496	49680	52000	4500	5040	5750
NET2	19,305	11040	12230	375	1141	1672
SWT2	14,844	11040	12230	675	1141	1672
MWT2	16,248	11040	12230	1144	1141	1672
AGLT2	21,855	11040	12230	1050	1141	1672
WT2	9,057	11040	12230	597	1141	1672
Facilities (tot)	112,805	104,880	113,150	8,341	10,205	14,110
Facilities (T2)	81,309	55,200	61,150	3,841	5,705	8,360

At 50% CPU capacity allows to simulate 800k events (full) or 12M events (fast) per day
At 50% CPU capacity allows to analyze 244M events (group) / 24400M events (user) per day

T1	U.S. ATLAS Tier-1 Center at BNL
NET2	North East Tier-2 Center at Boston University and Harvard University
SWT2	Southwest Tier-2 Center at University at Texas – Arlington and Oklahoma University
MWT2	Midwest Tier-2 Center at University of Chicago and Indiana University
AGLT2	ATLAS Great Lakes Tier-2 Center at University of Michigan and Michigan State University
WT2	Western Tier-2 Center at SLAC

ATLAS resource needs for 2010 – 2012

ATLAS total

CPU [kHS06]	Old 2010	Pledges 2010	New 2010	2011	2012
CERN	67	67	74	75	28
Tier-1	192	216	178	226	223
Tier-2	240	219	226	278	295
Disk [PB]	**Old 2010**	**Pledges 2010**	**New 2010**	**2011**	**2012**
CERN	3.9	3.9	4.7	7.0	7.5
Tier-1	21.9	22.3	22	25	27
Tier-2	20.9	21.3	24	38	44
Tape [PB]	**Old 2010**	**Pledges 2010**	**New 2010**	**2011**	**2012**
CERN	9.0	8.9	8.9	12.2	12.5
Tier-1	14.2	15.5	18	30	39

U.S. Share

CPU [kHS06]		Pledge 2010	New 2010	2011	2012
Tier-1		49.680	40.940	52.000	51.290
Tier-2		55.200	49.720	61.150	64.900
Disk [PB]					
Tier-1		5.037	5.060	5.750	6.210
Tier-2		5.700	5.280	8.360	9.680
Tape [PB]		3.266	4.140	6.900	8.970